1.	Record Nr.	UNINA9910847589903321
	Autore	Atanasova Pepa
	Titolo	Accountable and Explainable Methods for Complex Reasoning over Text / / by Pepa Atanasova
	Pubbl/distr/stampa	Cham : , : Springer Nature Switzerland : , : Imprint : Springer, , 2024
	ISBN	3-031-51518-8
	Edizione	[1st ed. 2024.]
	Descrizione fisica	1 online resource (208 pages)
	Disciplina	006.31
	Soggetti	Natural language processing (Computer science)
		Information storage and retrieval systems
		Machine learning
		Natural Language Processing (NLP)
		Information Storage and Retrieval
		Machine Learning
		Aprenentatge automàtic
		Tractament del llenguatge natural (Informàtica)
		Sistemes d'informació
		Llibres electrònics
	Lingua di pubblicazione	Inglese
	Formato	Materiale a stampa
	Livello bibliografico	Monografia
	Nota di contenuto	1. Executive Summary -- Part I: Accountability for Complex Reasoning Tasks over Text -- 2. Fact Checking with Insufficient Evidence -- 3. Generating Label Cohesive and Well-Formed Adversarial Claims -- Part II: Explainability for Complex Reasoning Tasks over Text -- 4. Generating Fact Checking Explanations -- 5. Generating Fluent Fact Checking Explanations with Unsupervised Post-Editing -- 6. Multi-Hop Fact Checking of Political Claims -- Part III: Diagnostic Explainability Methods -- 7. A Diagnostic Study of Explainability Techniques for Text Classification -- 8. Diagnostics-Guided Explanation Generation -- 9. Recent Developments on Accountability and Explainability for Complex Reasoning Tasks.
	Sommario/riassunto	This thesis presents research that expands the collective knowledge in the areas of accountability and transparency of machine learning (ML)

models developed for complex reasoning tasks over text. In particular, the presented results facilitate the analysis of the reasons behind the outputs of ML models and assist in detecting and correcting for potential harms. It presents two new methods for accountable ML models; advances the state of the art with methods generating textual explanations that are further improved to be fluent, easy to read, and to contain logically connected multi-chain arguments; and makes substantial contributions in the area of diagnostics for explainability approaches. All results are empirically tested on complex reasoning tasks over text, including fact checking, question answering, and natural language inference. This book is a revised version of the PhD dissertation written by the author to receive her PhD from the Faculty of Science, University of Copenhagen, Denmark. In 2023, it won the Informatics Europe Best Dissertation Award, granted to the most outstanding European PhD thesis in the field of computer science.